# การรู้จำอิริยาบถของมนุษย์โดยใช้ข้อมูลจากเซนเซอร์ของสมาร์ทโฟน สำหรับการยืนยันตัวตนผู้ใช้งาน

ณัฐพร โสมหิรัญ, รุ่งเรืองทรัพย์ สุขปรุง และ ผศ.ดร.วราภรณ์ วิยานนท์
หลักสูตรวิทยาศาสตรบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

## 1 Abstract

Smartphones are devices that help users to facilitate various functions anywhere anytime. Traditional ways to protect users' data on mobile devices with active user authentication approaches such as PINs, passwords, and fingerprint recognition. The passive and continuous user authentication will secure users' data that users do not need to manually verify their identity and repeat verification for a certain period of time.

Therefore, we have developed a passive and continuous user authentication application to reduce the risk of being attacked by any attempts to access the smartphone by using behavioral biometrics and machine learning methods. We developed the models for the application development consisting of 1) models for activity recognition and 2) models for user authentication. The data were collected from four positions of the user's body: right pocket, wrist, upper arm, and belt. To develop the models for activity recognition, there are two steps consisting of (1) feature extraction (2) training process using SVM, k-NN, Decision tree, and Bayesian networks. To develop the models for user authentication there are three steps consisting of (1) feature extraction (2) feature selection (3) training process using one-class SVM and one-class k-NN to verify user authentication.

From experimental results for activity recognition, the model from SVM can achieve excellent activity recognition performance with 99.48% on an upper arm. For the user authentication model, one-class SVM returns the magnificent results of 97.35% with sitting activity and pocket position. These models will be used to develop the passive and continuous user authentication application.

## 2 องค์ความรู้ที่เกี่ยวข้อง

**1. การเตรียมข้อมูล (Data pre-processing)**

- Noise removal [1]
- Data segmentation [2]
- Feature extraction [3]
- Feature selection [4]

**2. การประยุกต์ใช้การทำเหมืองข้อมูล (Data mining) เพื่อรู้จำอิริยาบถ**

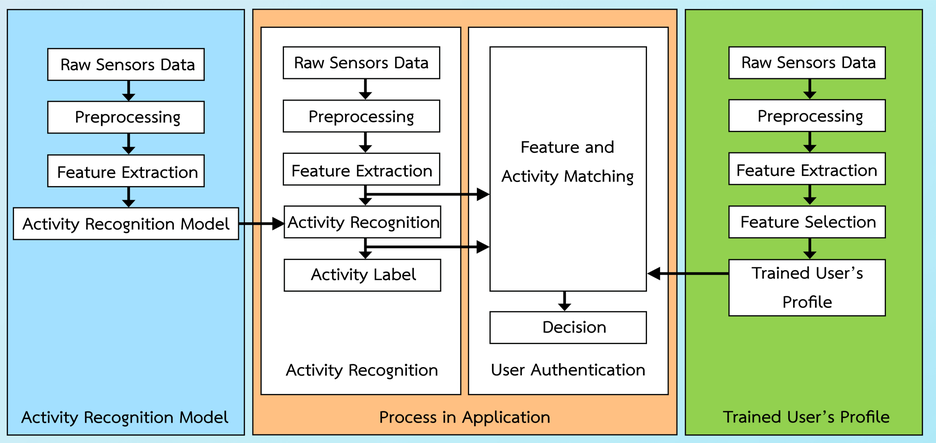
การทำเหมืองข้อมูล (Data mining) เพื่อจำแนกอิริยาบถต่างๆ ได้แก่ การเดิน การนั่ง การยืน การวิ่ง การเดินขึ้นบันได และการเดินลงบันได

**3. การยืนยันตัวตนผู้ใช้งานจากการรู้จำอิริยาบถ**

One-class Support Vector Machine [5]
เป็นอัลกอริทึมที่ปรับวิธีการมาจาก SVM เพื่อแก้ปัญหาการจำแนกประเภทของคลาสเดียว

One-Class k-Nearest Neighbors [5]
เปรียบเทียบข้อมูลคำนวณค่าระยะห่าง

## 3 ขั้นตอนการดำเนินงาน




ภาพแสดง system overview

### Preprocessing

กำจัดสัญญาณรบกวนด้วย Average smoothing method จากนั้นคำนวณและเพิ่มค่า magnitude ของทั้ง 3 เซนเซอร์ ทำให้แต่ละเซนเซอร์ประกอบด้วยค่าแกน x แกน y แกน z และ magnitude และแบ่งข้อมูลออกเป็นส่วนย่อยโดยใช้วิธี fixed-size window segmentation กำหนดขนาดเท่ากับ 5 วินาที ซึ่งเพียงพอต่อการรู้จำอิริยาบถ [2]

### Feature selection

ทำการเลือกคุณลักษณะด้วยวิธี ReliefF [7] โดยอัลกอริทึมจะให้คะแนนของแต่ละคุณลักษณะที่นำเข้าโดยคำนวณจากระยะห่าง Euclidean จากนั้นทำการเลือกคุณลักษณะ 20 ลำดับแรกมาใช้ในจำแนกโดยใช้ Gridsearch

### Dataset [6]

1 ไฟล์ข้อมูลแบ่งจำแนกตามผู้ใช้งานทั้งหมด 10 คน และตำแหน่งของสมาร์ทโฟนทั้งหมด 4 ตำแหน่ง คือ กระเป๋ากางเกงขวา ต้นแขนขวา ข้อมือขวา และเอว ข้อมูลภายในไฟล์ประกอบด้วย Activity label ค่าแกน x แกน y และแกน z ของเซนเซอร์วัดค่าความเร่ง เซนเซอร์วัดการหมุน และเซนเซอร์วัดสนามแม่เหล็ก

### Feature extraction

ทำการสกัดคุณลักษณะของแต่ละกลุ่มข้อมูล (Data segment) โดยคุณลักษณะที่เลือกพิจารณาได้แก่ max, min, variance, mean, skewness, kurtosis, peak

## 4 ผลการดำเนินงาน

1. ผลการประเมินประสิทธิภาพของการรู้จำอิริยาบท

ค่าพารามิเตอร์ที่ได้จากการใช้ Gridsearch ของแต่ละอัลกอริทึมมีดังนี้

- Support Vector Machine: kernel = linear, C = 1.0, gamma = 0.01
- Decision tree: max_depth = 5, min_samples_split = 7
- Bayesian networks: priors = None
- k-Nearest Neighbors: k = 1

ตารางเปรียบเทียบค่าความถูกต้องของอัลกอริทึมและตำแหน่ง

| Algorithm | Position | | | |
|---|---|---|---|---|
| | pocket | arm | belt | wrist |
| Decision Tree | **0.9846** | 0.9844 | 0.9742 | 0.9534 |
| Support Vector Machines | 0.9744 | **0.9948** | 0.9227 | **0.9689** |
| k-Nearest Neighbors | 0.8615 | 0.9948 | **0.9742** | 0.8446 |
| Bayesian networks | 0.9589 | 0.9430 | 0.8557 | 0.9637 |

2. ผลการประเมินประสิทธิภาพของการยืนยันตัวตน

ค่าพารามิเตอร์ที่ได้จากการใช้ Gridsearch ของแต่ละอัลกอริทึมมีดังนี้

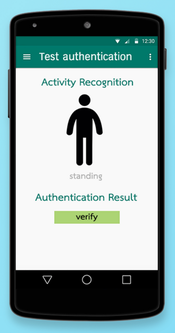
- One-class Support Vector Machine: kernel = rbf, gamma = 0.001, nu = 0.025
- One-class k-Nearest Neighbors: k = 7

ตารางเปรียบเทียบค่าความถูกต้อง อัตราการปฏิเสธผิดพลาดและอัตราการอนุญาตผิดพลาดของอัลกอริทึมและตำแหน่ง

| Activity | Parameter | Algorithm | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | SVM | | | | k-NN | | | |
| | | pocket | arm | belt | wrist | pocket | arm | belt | wrist |
| walking | FRR | 13.25 % | 10.13 % | 16.00 % | **7.36 %** | 25.71 % | 24.29 % | 25.71 % | 28.57 % |
| | FAR | 5.36 % | 12.41 % | **3.90 %** | 13.20% | 11.49 % | 12.57 % | 4.92 % | 15.52 % |
| | Accuracy | 94.39 % | 87.32 % | **95.46 %** | 86.98 % | 87.76 % | 86.52 % | 93.98 % | 83.54 % |
| running | FRR | 15.11 % | 14.25 % | 17.25 % | **13.47 %** | 30.00 % | 20.00 % | 30.00 % | 32.86 % |
| | FAR | **2.95 %** | 7.62 % | 8.15 % | 6.76 % | 9.11 % | 6.00 % | 5.17 % | 4.48 % |
| | Accuracy | **96.52 %** | 91.87 % | 91.25 % | 92.43 % | 90.43 % | 93.04 % | 93.58 % | 94.47 % |
| sitting | FRR | 19.25 % | 18.00 % | **16.60 %** | 21.11 % | 28.43 % | 31.43 % | 35.71 % | 41.43 % |
| | FAR | **2.03 %** | 3.33 % | 4.00 % | 3.65 % | 2.73 % | 4.98 % | 2.44 % | 5.65 % |
| | Accuracy | **97.35 %** | 95.90 % | 95.35 % | 95.63 % | 96.21 % | 94.01 % | 96.40 % | 93.14 % |
| standing | FRR | 17.00 % | 17.00 % | **16.25 %** | 26.00 % | 28.57 % | 27.00 % | 32.86 % | 25.71 % |
| | FAR | 3.65 % | 7.14 % | **3.14 %** | 14.35 % | 7.56 % | 15.49 % | 9.68 % | 13.75 % |
| | Accuracy | 95.93 % | 92.36 % | **96.46 %** | 85.13 % | 91.34 % | 83.54 % | 89.16 % | 87.12 % |
| upstairs | FRR | 15.50 % | 11.11 % | 19.25 % | **7.50 %** | 25.71 % | 25.71 % | 45.71 % | 27.14 % |
| | FAR | 17.42 % | 40.70 % | 5.52 % | 39.14 % | 20.44 % | 23.30 % | **5.24 %** | 35.27 % |
| | Accuracy | 82.64 % | 60.18 % | **94.05 %** | 61.82 % | 78.79 % | 75.99 % | 93.88 % | 64.25 % |
| downstairs | FRR | **19.00 %** | 19.50 % | 22.30 % | 23.61 % | 30.00 % | 26.14 % | 35.71 % | 37.14 % |
| | FAR | 17.90 % | 27.55 % | 19.90 % | 22.57 % | 27.90 % | 27.17 % | **7.84 %** | 39.56 % |
| | Accuracy | 82.03 % | 72.67 % | **94.63 %** | 77.39 % | 71.40 % | 71.96 % | 90.90 % | 60.06 % |

FRR = False Rejection Rate, FAR = False Acceptance Rate

## 5 สรุปผลการดำเนินงาน




จากการทดสอบพบว่าโมเดลการรู้จำอิริยาบถจากอัลกอริทึม Support Vector Machine ให้ประสิทธิภาพดีกว่าโมเดลจากอัลกอริทึมอื่นๆ ที่นำมาทดสอบ โดยมีค่าความถูกต้อง 99.48% สำหรับบริเวณต้นแขน

ส่วนโมเดลการยืนยันตัวตนนั้น อัลกอริทึมที่ให้ประสิทธิภาพที่ดีคือ One-class SVM โดยมีค่าความถูกต้อง 97.35% สำหรับอิริยาบถการนั่งและมีตำแหน่งของสมาร์ทโฟนอยู่ที่บริเวณกระเป๋ากางเกง

จากนั้นนำวิธีการรู้จำอิริยาบถและการยืนยันตัวตนผู้ใช้งานที่ได้ศึกษามาพัฒนาแอปพลิเคชันต้นแบบแอนดรอยด์สำหรับการยืนยันตัวตนผู้ใช้งาน

## 6 บรรณานุกรม

[1] X. Su, H. Tong, and P. Ji, Activity recognition with smartphone sensors, Tsinghua Science and Technology, vol. 19, no. 3, pp. 235–249, 2014

[2] L. Wang, T. Gu, X. Tao, and J. Lu, A hierarchical approach to real-time activity recognition in body sensor networks, Pervasive and Mobile Computing, vol. 8, no. 1, pp. 115- 130, 2012.

[3] M. Khan, S. I. Ahamed, M. Rahman, and R. O. Smith, A feature extraction method for realtime human activity recognition on cell phones, in Proc. 3rd Int. Symposium on Quality of Life Technology Conf., Toroto, Canada, 2007.

[4] An Introduction to Feature Selection. Retrieved February 15, 2019, from http://www.datagenetics.com/blog/september32013

[5] S. S. Khan and M. G. Madden, "One-class classification: taxonomy of study and review of techniques," Knowledge Engineering Review, vol. 29, no. 3, pp. 345–374, 2014.

[6] Physical activity recognition dataset, Sensors activity dataset. Retrieved December 21, 2018, from https://www.utwente.nl/en/eemcs/ps/research/dataset.

[7] R. J. Urbanowicza, M. Meeker, W. L. Cava, R. S. Olson, andJ. H. Moore, "Relief-based feature selection: introduction andreview," Journal of Biomedical Informatics, vol. 85, pp. 189–203, 2018